3D MR Image Restoration by Combining Local Genetic Algorithm with Adaptive pre-conditioning

Tianzi Jiang [1,2,*] *Member, IEEE*, and Frithjof Kruggel [1]
[1] Max-Planck Institute of Cognitive Neuroscience
Stephanstrasse 1, D-04103, Leipzig, Germany
Email: {jiang, kruggel}@cns.mpg.de
[2] National Laboratory of Pattern Recognition, Institute of Automation
Chinese Academy of Sciences, Beijing 100080,P.R. China
Email: jiangtz@sycamore.ia.ac.cn

Abstract

In this paper, we propose a novel efficient method by incorporating a local genetic algorithm and a new pre-conditioning technique into Markov random field model for image restoration. The role of genetic algorithm is to improve the quality of restoration and the pre-conditioning technique aims at accelerating the convergence. The remarkable advantage of our approach is that restoring corrupted images and preserving the shape transitions in the restored results have been orchestrated very well. The experiments on 3D MR image show that our method work very well.

1 Introduction

Image restoration has been widely investigated in the field of image processing. Efficient restoration has proven to be very useful for many image processing applications. Markov random fields is one of the most powerful approaches to image restoration. In [2], Geman and Geman proposed a Bayesian framework for image restoration using Markov Random Fields. The use of MRF enables the integration of some general priors. Since one pixel value of an image is not independent but has spatial dependencies on the values of its neighbors. This contextual prior knowledge must be enclosed in our model. MRFs give an appropriate description of the interactions between neighboring pixels. Because there are many discontinuities in images, especially in the areas near edges, we need to control the interaction between neighboring pixels to avoid over-smoothed solutions. For image denoising, the most common problem is that some interesting structures in the image will be removed from original image during noise suppression. Such interesting structures in an image often correspond to discontinuities in the image. The major motivation of our work is to develop an algorithm which can both reduce the degradations and preserve discontinuities.

In this paper, we consider the 3D MR image or volume denoising problem in the framework of MRFs. This problem is an essential preprocessing step for many applications of medical image analysis. We introduce a novel hybrid search algorithm for this problem. That is, we incorporate a local genetic algorithm and a novel pre-conditioning technique into Markov random field model for image restoration. The experiments on 3D MR image show that our method work very well. Comparing with the Wiener filter, it is much more effective and powerful.

The paper is organized as follows. Section 2 devotes to Markov Random Field formulation for image restoration. In Section, we give a brief description of of our local genetic algorithm. Section 4 is about a novel pre-conditioning technique. A preliminary experimental result on 3D MR image is given in Section 5. The final section is a conclusion.

2 MRF Formulation for Image Restoration

In this section, we present MRF formulation of image restoration. Let $\mathbf{F} = \{\mathbf{F}_{ij} | (i,j) \in \mathcal{S}\}$ denote a family of random variables over the lattice $\mathcal{S}$, in which each random variable $\mathbf{F}_{ij}$ takes a value $\mathbf{f}_{ij}$ in G. In what follows, G will denote the set of all the possible gray levels for a pixel. Let Ω be the state space of the random field $\mathbf{F}$, which is said to be a Markov Random Field on $\mathcal{S}$ with respect to a neighborhood system $\mathcal{N}$ if and only if $\forall \mathbf{f} \in \Omega$, $P(\mathbf{F} = \mathbf{f}) > 0$, and

$$\begin{aligned} & P(\mathbf{F}_{ij} = \mathbf{f}_{ij} | \mathbf{F}_{kl} = \mathbf{f}_{kl}, (k,l) \in \mathcal{S} - \{(i,j)\}) \\ = \; & P(\mathbf{F}_{ij} = \mathbf{f}_{ij} | \mathbf{F}_{kl} = \mathbf{f}_{kl}, (k,l) \in \mathcal{N}_{(i,j)}\} \qquad (1) \end{aligned}$$

*From April 1, 2000, adress correspondence to Tianzi Jiang, School of Computer Science, The Queen's University of Belfast, Belfast BT7 1NN, UK, email: T.Jiang@qub.ac.uk

0-7695-0750-6/00 $10.00 © 2000 IEEE

Form the Hammersley-Clifford theorem, the joint probability distribution of an MRF can be written as a Gibbs distribution $P(\mathbf{F}=\mathbf{f})=\frac{1}{Z}\exp(-U(\mathbf{f}))$, $\forall \mathbf{F}\in\Omega$ where $Z=\sum_{\mathbf{f}\in\Omega}\exp(-U(\mathbf{f}))$ is a normalization constant and $U(\mathbf{f})$ is the energy function defined as $U(\mathbf{f})=\sum_{c\in C}V_c(\mathbf{f}_{ij},(i,j)\in c)$, where $C=C_1\cup C_2\cup C_3\cdots$ is the set of cliques and V_c is the *clique potential* associated with the clique c.

For image restoration, a coupled MRF model may be constructed by defining an MRF, $\mathbf{F}$, for image, and a boolean MRF, $\mathbf{L}$, for the line processes. The prior probability can be defined as follows.

$$P(\mathbf{F}=\mathbf{f},\mathbf{L}=\mathbf{l})=\frac{1}{Z}\exp(-U(\mathbf{f},\mathbf{l}))/T \qquad (2)$$

where $U(\mathbf{f},\mathbf{l}) = \sum_{c\in C}V_c(\mathbf{f},\mathbf{l})$, and $Z = \sum_{(\mathbf{f},\mathbf{l})}\exp(-U(\mathbf{f},\mathbf{l}))$, The parameter T is assumed to be 1 for simplicity. The prior energy $U(\mathbf{f},\mathbf{l})$is a smoothness constraint on the image and the line process.

Let $\mathbf{d}$ denote the observed image contaminated by an additional noise $\mathbf{N}$. The observation model can be described as $\mathbf{D}=H(\mathbf{F})+\mathbf{N}$, where $\mathbf{N}$ is zero mean white Gaussian radom field with variance σ for each variable in $\mathbf{N}$. According to the Bayes rule, we have

$$\begin{aligned} & P(\mathbf{D}=\mathbf{d}|\mathbf{F}=\mathbf{f},\mathbf{L}=\mathbf{l}) \\ \propto \ & P(\mathbf{F}=\mathbf{f},\mathbf{L}=\mathbf{l}|\mathbf{D}=\mathbf{d})P(\mathbf{D}=\mathbf{d}) \end{aligned} \qquad (3)$$

Then MAP estimate of (3) is found by

$$y^*=\arg\max_{\mathbf{d}\in\Omega}P(\mathbf{F}=\mathbf{f},\mathbf{L}=\mathbf{l}|\mathbf{D}=\mathbf{d})P(\mathbf{D}=\mathbf{d}) \quad (4)$$

which can be obtained by minimizing the following energy function

$$U(\mathbf{f},\mathbf{l}|\mathbf{d})=\frac{1}{2\sigma^2}\|\mathbf{d}-H(\mathbf{f})\|^2+\sum_c V_c(\mathbf{f}|\mathbf{l})+\sum_c V_c(\mathbf{l}) \qquad (5)$$

This is a non-convex function with a hybrid of real and boolean variables. If we denote the observation model by $\mathbf{d}=\mathbf{K}_d(\mathbf{f})+\mathbf{n}$, where $\mathbf{d}\in R^{n^2\times 1}$ is the data vector, the operator $\mathbf{K}_d$ is the stiff matrix resulted from the discretion, and $\mathbf{n}$ is zero mean white Gaussian radom field with variance σ for each variable in $\mathbf{n}$, we can rewrite the energy function $U(\mathbf{f},\mathbf{l}|\mathbf{d})$ as follows.

$$U(\mathbf{f},\mathbf{l}|\mathbf{d})=U(\mathbf{f}|\mathbf{l},\mathbf{d})+U(\mathbf{l}), \qquad (6)$$

$$U(\mathbf{f}|\mathbf{l},\mathbf{d})=\frac{1}{2\sigma^2}\|\mathbf{K}_d(\mathbf{f})-\mathbf{d}\|^2+U(\mathbf{f}|\mathbf{l}), \qquad (7)$$

$$U(\mathbf{f}|\mathbf{l})=\sum_c V_c(\mathbf{f}|\mathbf{l}), \quad \text{and } U(\mathbf{l})=\sum_c V_c(\mathbf{l}),$$

Extending the line processing to 3D, *the line-sites* edge elements p_{ijk}, q_{ijk}, and r_{ijk} representing small planar elements in the i-, j-, and k- directions respectively, so that for examples $p_{ijk}=1$ indicates an active edge between voxels f_{ijk} $f_{i+1,j,k}$. Inactive elements take the value -1. For the four connected neighborhood system, the conditional clique potential of the real variables $\mathbf{f}$ given the line process, $\mathbf{l}$, $U(\mathbf{f}|\mathbf{l})$, can be formulated as follows.

$$\begin{aligned} U(\mathbf{f}|\mathbf{l}) &= \frac{\nu}{2}\sum_{i,j,k}\big[(1-p_{i,j,k})(f_{i+1,j,k}-f_{i,j,k})^2 \\ &+ (1-q_{i,j,k})(f_{i,j,k}-f_{i,j+1,k})^2 \\ &+ (1-r_{i,j,k})(f_{i,j,k}-f_{i,j,k+1})^2\big] \end{aligned} \qquad (8)$$

where $\mathbf{l}=[\mathbf{p},\mathbf{q},\mathbf{r}]$. Then the MAP estimate problem can be cast as the following optimization problem:

$$\min_{\mathbf{f},\mathbf{l}}U(\mathbf{f},\mathbf{l}|\mathbf{d})=\min_{\mathbf{l}}\left\{U(\mathbf{l})+\min_{\mathbf{f}}U(\mathbf{f}|\mathbf{l},\mathbf{d})\right\} \qquad (9)$$

Define the function

$$E(\mathbf{l})=U(\mathbf{l})+\min_{\mathbf{f}}U(\mathbf{f}|\mathbf{l},\mathbf{d}) \qquad (10)$$

Then the original non-convex optimization problem is transformed to a hybrid of the combination optimization problem (10) and the continuous optimization problem: $\min_{\mathbf{f}}U(\mathbf{f}|\mathbf{l},\mathbf{d})$. Once it is understood that image restoration is such an optimization problem, the use of any technique for tackling optimization problem suggests itself. Previous works suggested that simulated annealing was used for the binary line process and an analog network to solve the continuous optimization problem [1]. For 2D image restoration, Lai *et al* [5] proposed a modified genetic algorithm for alternative to simulated annealing. In this paper, we propose a kind of local genetic algorithm to solve this problem. This algorithm has excellent global and local search performance.

3 Local Genetic Algorithm

Genetic algorithms (GA) are capable of searching for global optimum in functions which cause difficulty for gradient based methods [3]. Principal advantages of GA are domain independence, non-linearity and robustness. However, conventional GA has a very poor local performance because of the random search of GA. In this paper, several remarkable features are added and some important extensions are also made to improve the performance of the conventional GA. The implementation details of these features are described as follows.

(1) Hybrid Algorithm Since GA has poor local search performance and conventional local search methods have remarkable ability in finding local optima, we propose a hybrid algorithm, which combines GA with a modified local search procedure. In the hybrid algorithm, Hooke-Jeeves local search procedure [4], a simple but efficient local search method, is applied to each new solution generated by the genetic operations (i.e., selection, crossover, and mutation) to maximize its fitness value E. The local search is also applied to elite solutions inherited from the previous populations. Some parameters are introduced into the algorithm to control the local search. The parameter *Step* is used to control the initial step of the search. The parameter ε is used to control the distance from the solution to its nearest local optimum.

(2) Elitist Strategy Two sets of solutions are stored in our algorithm: a current population and an elite set. After genetic operation and local search, the current population is replaced with the improved population and the elite set is updated from the new population. A local search procedure is applied to solutions in the elite set. By preserving good solutions, we can avoid losing some excellent solutions. It has been shown that only after one adopts the elitist strategy, the GA can mathematically converge to a global optimum [3]. In our algorithm, the parameter *EliteRate* is used to control the number of the elite.

(3) Selection Operation and Crossover Operation In our algorithm, we adopt the roulette wheel selection mechanism. The selection probability $P(x)$ in the current population Φ can be expressed as $P(x) = \frac{f^{\lambda}(x)}{\sum_{x\in\Phi} f^{\lambda}(x)}$, where $f(x)$ is the fitness value of solution x and λ is the parameter to control the scaling procedure of selection operation. Different values of λ can make GA result in different solutions: when the value of λ is large, the selection mechanism is strong and competition in the group is intense, some outstanding solutions in the population have greater chance to survive. However, the GA is easily trapped in a local optimum because one or two outstanding solutions in the group will be dominant in the population rapidly and the evolution will stop at a suboptimal solution. Thus in the beginning of GA, we give λ a small value to limit the competition, and we increase λ in subsequent generations in order to stimulate the evolution.

4 Adaptive Pre-conditioning

As mentioned in Section 2, the image restoration can be cast as a hybrid of the combination optimization problem (10) and the following continuous optimization problem:

$$\min_{\mathbf{f}} U(\mathbf{f}|\mathbf{l}, \mathbf{d}). \quad (11)$$

For each line process configuration $\mathbf{l}$ visited by the LGA, we first need to solve a convex and quadratic optimization (11). This is equivalent to solve a linear system with an associated symmetric positive definite matrix (SPD). From (7) and (8), we can obtain the associated SPD matrix $\mathbf{K}$ for the convex quadratic function $U(\mathbf{f}|\mathbf{l}, \mathbf{d})$ as follows.

$$\mathbf{K} = \frac{1}{2\sigma^2}\mathbf{K}_d + \mathbf{K}_s \quad (12)$$

where $\mathbf{K}_d \in R^{N\times N}(N = n^2)$ and $\mathbf{K}_s \in R^{N\times N}$ is the matrix from the quadratic smoothness given in (8).

To solve the above linear system, various preconditioned iterative methods were proposed. In this paper, we propose an adaptive filtering technique for this question. We first give a brief review of tangential frequency filtering decompositions(TFFD), which is basis of our algorithm.

Definition 4.1 *The tangential frequency filtering decompositions(TFFD) of a block-tridiagonal matrix $K \in R^{n\times n}$*

$$K = \begin{pmatrix} D_1 & L_{1,2} & 0 & \dots & 0 \\ L_{2,1} & D_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & D_{n-1} & L_{n-1,n} \\ 0 & \dots & 0 & L_{n,n-1} & D_n \end{pmatrix} \quad (13)$$

for a given test vector $t = (e_1^T, \cdots, e_n^T)^T, t \in R^n$, is defined by

$$M = (L+T)T^{-1}(U+T).$$

L is the lower block-triangular part of K, U is the upper block-triangular part of K and T is a block-diagonal matrix $T = blockdiag(T_1, \cdots, T_n)$ with blocks $T_1 = D_1$ and for $i > 1$

$$T_i = D_i + \Theta_{i,i-1}T_{i-1}\Theta_{i-1,i} - \Theta_{i,i-1}L_{i-1,i} - L_{i,i-1}\Theta_{i-1,i}.$$

The choice of the transfer matrices $\Theta_{i,j}$ is restricted to those which fulfill the filter conditions

$$\begin{aligned}[\Theta_{i,i-1}T_{i-1}\Theta_{i-1,i} &- \Theta_{i,i-1}L_{i-1,i} - L_{i,i-1}\Theta_{i-1,i}]e_i \\ &= L_{i,i-1}T_{i-1}^{-1}L_{i-1,i}e_i, \quad (14)\end{aligned}$$

for the blocks e_i of the test vector t.

If in (13), $L_{ij} = L = L^T$, $D_i = D = D^T$, $D, L \in R^{m\times m}$, then the correct test vectors

$$t^{(i)} = ({e^{(i)}}^T, \cdots, {e^{(i)}}^T) \in R^{n\cdot m} \quad (15)$$

consists for theses matrices of the eigenvectors $e^{(i)}$ of the generalized eigenvalue problem

$$\rho^{(i)}De^{(i)} = Le^{(i)}, \quad |\rho^{(i)}| \le \frac{1}{2}, \quad i = 1, 2, \cdots, m.$$

Algorithm 4.1 *For $i = 1, 2, \cdots, q$, $m \in N$, an iterative scheme to approximate to the solution of the linear system*

$$Ku = f, \quad K \in R^{k \times k}, \quad u, f \in R^k,$$

based on a sequence $\{M^{(i)}\}_{i=1}^{q}$ of FFDs or TFFDs is given by

$$u^{(m+\frac{i}{q})} = u^{(m+\frac{i-1}{q})} + {M^{(i)}}^{-1}(f - Ku^{(m+\frac{i-1}{q})}),$$

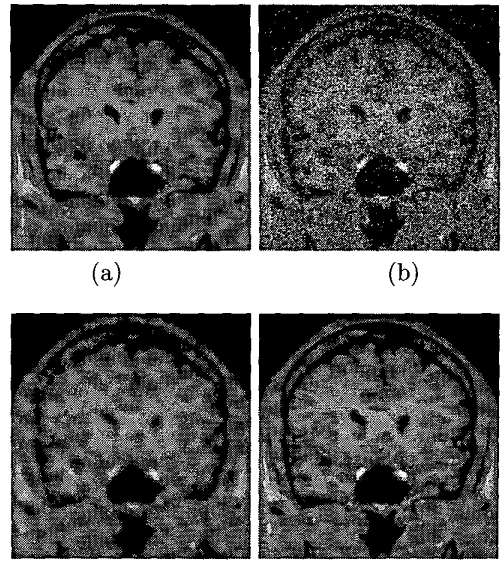
Theorem 4.1 *Let a SPD system matrix K of the form (13) with $L_{ij} = L$, $D_i = D$ and a suitable sequence of the test vectors be given. Then the Algorithm 4.1 based on the corresponding sequence of TFFDs or FFDs $M^{(i)}$, $i = 1, 2, \cdots, q$, convergences independently of the number of unknowns $\|S_\Pi\|_K < \sigma < 1$, where $S_\Pi = \prod_{i=1}^{q} S^{(i)}$, $S^{(i)} = I - {M^{(i)}}^{-1}K$.*

5 Experimental Results

We illustrate the efficacy and power of our approach using 3D MR image corrupted by Gaussian white noise with mean of zero and variance of 0.05. Comparison of our algorithm with the result from Wiener adaptive filtering is also made. We choose these this filtering technique only because it is the most typical as well as widely used ones.

Figure 1(a) is raw MRI volume. Figure 1(b) is test image obtained by adding Gaussian white noise with mean of zero and variance of 0.05 to the raw volume. Figure 1(c) is the result obtained with 2D wiener filtering which is performed slice by slice. Figure 1(d) is the result obtained with our approach.

Our experimental result demonstrate that our approach is much more effective and powerful in noise reduction.



(a) (b) (c) (d)

Figure 1: (a) Original 3D 165×160×3 MRI image; (b) Image corrupted with Gaussian noise, the variance is about 0.05; (c) Result of restoration by 2D Wiener filter with 3× neighborhood; (d) Result obtained through combining MRF and genetic algorithm.

6 Conclusion

In this paper, We have proposed a novel and efficient method to improve the denoising of 3D MR image, which is a crucial preprocessing step for the further analysis. The advantage of our approach is that restoring corrupted image and preserving the shape transitions in the restored results have been orchestrated very well. Experiments illustrate that our method is much more effective and powerful than the Wiener technique, a typical and widely used technique. We are also undergoing research into the comparison of our method developed in this paper with the pixon-based methods for some typical problem in medical image analysis.

References

[1] L. Bedini and A. Tonazzini, "Image restoration preserving discontinuities: the Bayesian approach and neural networks", Image and Vision Computing, vol. 10, no. 2, 1992, 108-118.

[2] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 6, no. 6, pp. 721-741, 1984.

[3] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*, Reading, MA: Addison-Wesley, 1989.

[4] J. P. Ignizio, *Goal Programming and Extensions*, Lexington, MA: Lexington Books, 1976.

[5] S. H. Lai, B. C. Vemuri, "Efficient hybrid search for visual reconstruction problems", Image And Vision Computing, vol. 17, no. 1, pp. 37-49 1999.

[6] C. Wagner and G. Wittum, "Tangential Frequency Filtering Decompositions for Symmetric Matrices", Numer. Math., 78:1(1997), 119-142.